基于用户潜在兴趣的知识感知传播推荐算法 *

张　波，赵　鹏†，张金金，曾昭菊，肖栩豪
(中国人民解放军火箭军工程大学 作战保障学院，西安 710025)

摘　要：知识图谱引入推荐系统可以利用知识图谱实体之间的语义关系学习用户及项目表示。基于嵌入传播的方法利用知识图谱的图结构学习相关特征，但随着传播范围增加，多跳实体间的语义相关性减小。为有效提升推荐语义表达能力并提高推荐准确度，文章提出基于用户潜在兴趣的知识感知传播推荐模型，该模型采用异构传播方式传播项目关联知识并迭代学习用户的潜在兴趣，以此增强模型对用户与项目的表示能力。具体的，首先图嵌入层生成用户与项目的初始化表示，随后在异构传播层中采用知识感知注意力机制区分同一层中实体之间的重要性，更精确的生成目标实体的表示。随后通过用户潜在兴趣传播学习用户的高阶潜在兴趣，增强多跳实体语义相关性。最后在预测层中使用信息衰减因子区分不同传播层次的重要性，生成用户及项目的最终表示。实验表明，该模型在 Last。FM 与 Book-Crossing 两个公开数据集上 AUC 值相较于最先进的基线提升了 2.25%与 4.71%，F1 值分别提升 3.05%和 1.20%，Recall@K 值均优于对比的基线模型，文章提出的模型能有效提高推荐准确度。

关键词：推荐系统；知识图谱；注意力机制；异构传播
中图分类号：TP391　　doi: 10.19734/j.issn.1001-3695.2022.03.0067

Knowledge-aware propagation recommendation algorithm based on user's potential interest

Zhang Bo, Zhao Peng†, Zhang Jinjin, Zeng Zhaoju, Xiao Xuhao
(*College of Operational Support*, *Rocket Force University of Engineering*, *Xi'an* 710025, *China*)

Abstract: Applying knowledge graph to recommendation system can make use of semantic relations between entities of knowledge graph to learn user and item representation. The embedding propagation method uses the graph structure of the knowledge graph to learn relevant features, but the semantic dependency between multi-hop entities decreases as the propagation range increases. In order to effectively improve the semantic expression ability of recommendation and improve the accuracy of recommendation, this paper proposes a knowledge-aware propagation recommendation algorithm based on users' potential interests. The model adopts heterogeneous propagation method to disseminate item relevant knowledge and iteratively learn users' potential interests, so as to enhance the representation ability of the model to users and items. Specifically, first, graph embedding layer generate initialize representation of users and items, and in the heterogeneous propagation layer, the knowledge-aware attention mechanism can distinguish the importance of entities in the same layer, so the model can generate the representation of target entities more accurately. Then the user's potential interest propagation can effectively learn the user's higher-order potential interest and enhance the semantic relevance of multi-hop entities. Finally, information decay factor is used in the prediction layer to distinguish the importance of different communication levels and generate the final representation of users and items. Experiments show that the AUC value of the model on the Last. FM and Book-Crossing increases by 2.25% and 4.71% compared with the most advanced baseline, and the F1 value increases by 3.05% and 1.20% respectively, and the Recall@K value is superior to the comparison baseline model. The proposed model can effectively improve the accuracy of recommendation.

Key words: recommended systems; knowledge graph; attention mechanisms; heterogeneous propagation

0　引言

计算机科学与互联网的发展使电子商务、数字媒体等各类数字化平台快速崛起，在提高人们生活水平的同时也使人们面临数据过载问题[1]。推荐系统可以在用户需求未知的情况下根据用户交互行为推荐用户可能喜欢的项目，从而缓解数据过载问题。协同过滤推荐算法[2~5]是推荐系统研究领域的重点，其基本思想体现了用户偏好相似与用户行为相似的关联特性，但当用户交互信息较少甚至无交互信息时就会产生数据稀疏问题[6]。为解决数据稀疏问题，研究人员开始关注辅助信息，利用辅助信息更好的学习项目相关特征并生成项目表示。

知识图谱推荐将知识图谱[7]引入到协同过滤推荐中，利用丰富的外部项目语义知识提高模型表达能力。知识图谱是一种结构化的辅助信息，由(头实体，关系，尾实体)三元组形式构成，其包含大量的项目背景与属性，这些信息由实体与关系表示。项目由属性相关联，因此在学习项目特征的过程中通过知识图谱的网络结构可以获取项目之间的语义相关性进而提升模型项目表示与物品描述能力。相较于一般协同过滤推荐，知识感知推荐可以使推荐结果更准确且具有一定的可解释性。

知识图谱推荐系统分为三种，分别是基于嵌入的方法、基于路径的方法以及基于嵌入传播的方法。其中基于嵌入传播的方法[8]将知识图谱语义嵌入与基于知识图谱路径的推荐

收稿日期：2022-03-01；**修回日期：**2022-04-15　**基金项目：**军队科研资助项目

作者简介：张波(1997-)，男，山西阳泉人，硕士研究生，主要研究方向为推荐系统；赵鹏(1979-)，男(通信作者)，河南孟县人，副教授，硕导，博士，主要研究方向为智能信息处理与仿真(zpxhh@163.com)；张金金(1992-)，女，陕西西安人，讲师，博士，主要研究方向为推荐系统、数据挖掘；曾昭菊(1999-)，女（壮族），广西融安人，硕士研究生，主要研究方向为推荐系统；肖栩豪(1997-)，男，湖北英山人，硕士，主要研究方向为智能信息处理.

方法相结合，增强了知识图谱中具有多跳邻居的实体表示。但该方法在高阶传播过程中语义知识相关性会随着传播层的增加而减小。例如 KGAT 模型[9]在使用图神经网络迭代学习实体嵌入表示时，通常迭代次数不超过 3 次。使用异构传播的 HKIPN 模型[10]中传播深度为 2、3 层时推荐效果最佳。然而，高阶实体的特征也具有一定的参考价值，当前知识感知推荐算法并不能有效利用高阶信息。为解决以上问题，本文提出一种基于用户潜在兴趣的知识感知传播推荐模型。该模型增强用户高阶潜在兴趣表示，解决异构传播高阶传播过程中存在的知识语义关联性降低的问题。

本文的主要贡献如下：

a)采用知识感知注意力机制区分异构传播中每层实体的邻居重要性，通过用户潜在兴趣传播，获得用户高阶潜在兴趣，增强用户及项目表示。

b)使用信息衰减因子区分传播层的重要性，增加模型的表达能力。

c)实验结果显示该模型在知识感知推荐系统主流数据集 Last.FM 与 Book-Crossing 上进行的 CTR 与 TOP-K 任务中都有良好的表现。

1　相关工作

近年来，知识图谱广泛应用于知识图谱补全[11]、问答系统[12]以及词嵌入[13]等研究领域，与此同时在推荐领域同样受到极大的关注。知识图谱推荐算法可以分为以下三类：

a)基于嵌入的方法(embedding-based method)是指利用图嵌入技术对知识图谱中的实体与关系进行嵌入表示。Zhang 等人[14]提出 CKE(Collaborative knowledge base embedding)模型，在协同过滤推荐的基础上将项目的知识图谱中相关属性与知识结构化嵌入表示，以此对项目进行扩充表示。He 等人[15]将用户个人信息融合到知识图谱中，生成异质信息网络并进行图嵌入表示，该模型充分利用用户与项目的辅助信息改善推荐效果。虽然基于嵌入的方法一定程度上可以有效的提升推荐性能，但严重依赖图嵌入算法，并且缺乏对知识图谱的高阶建模能力。

b)基于路径的方法(path-based mothod)利用知识图谱结构化的特征，挖掘图谱中实体之间的关系。在推荐系统算法中常常将用户-项目交互信息与知识图谱相融合，因此可以使用元路径挖掘方法学习实体之间的关系。Hu 等人[16]提出的基于元路径的推荐算法认为知识图谱也是特殊的异质信息网络，因此在知识图谱上以元路径的潜在特征提取方式进行连接，并生成不同类型的连接表示。Wang 等人[17]提出了一种知识感知路径递归网络(KPRN)，该模型通过实体嵌入和关系嵌入来构造提取的路径序列。使用 LSTM 层对各条路径进行编码，并且通过全连接层预测每个路径上用户对项目的偏好，最终加权聚合得到用户偏好估计。基于路径的方法需要一定的专业领域知识，并且随着路径数量的增加，处理耗时耗力且难以优化到理想效果，因此实用性较低。

c)基于嵌入传播的方法(embedding propagation-based methods)可以有效地利用知识图谱中的高阶信息，因此极大的提高了推荐准确度。该方法在知识图谱上以迭代传播关联知识的方式挖掘辅助信息，利用获取的辅助信息开展推荐任务以此提高推荐效果。Wang 等人[18]提出的 RippleNet(Multiple Ripples Networks)模型在知识图谱中迭代挖掘用户的潜在偏好，以用户的历史交互为种子，以“涟漪”的形式传播以此获得用户的层次兴趣。Wang 等人[19]提出基于图卷积网络的 KGCN(Knowledge Graph Convolutional Networks)模型，将知识图谱转换为用户特定条件下的加权图，并区分图中实体节点的重要性，迭代聚合邻居信息获得项目嵌入表示，该模型将图卷积网络应用到知识图谱推荐算法中并有效提高了推荐性能。随后基于 KGCN 模型提出 KGNNLS(Knowledge-aware Graph Neural Networks with Label Smoothness regularization)模型[20]，该模型使用标签平滑法将边的正则化等价于在知识图谱上进行标签传播。以上两种方法表明聚合邻居信息生成相关表示可以有效地提高推荐性能。Wang 等人[9]提出了 KGAT(Knowledge Graph Attention Network)模型，使用 TransR 算法学习知识图谱的嵌入表示，使用图注意力网络递归地从节点的邻居传播嵌入表示，以此增强节点的嵌入表示并生成权重区分邻居重要性，该模型进一步验证了嵌入传播对高阶关系探索的有效性。Chen 等人[10]提出 HKIPN(Hierarchical Knowledge and Interest Propagation Network)模型，使用协同知识图异构传播方法，基于多层注意力机制生成不同权重的高阶邻居信息，以此获得用户项目表示。Xu[21]等人提出了协同知识感知图注意网络(Collaborative Knowledge-Aware Graph Attention Network，CKGAT)，基于异构传播策略，CKGAT 模型使用知识感知图注意力网络提取多跳波纹集中实体的拓扑邻近结构，然后学习高阶实体表示，从而生成精细的波纹集嵌入。

基于嵌入传播的推荐可以有效结合基于嵌入的推荐方法与基于路径的推荐方法二者的优点，利用用户、项目的低维向进行交互计算，通过传播的方式实现多跳关系的挖掘。现有工作的主流思想是使用图神经网络探索知识图谱的多跳关系，如基于图神经网络的 CKGAT 模型。但现实中用户项目交互类型广泛使得高阶交互潜在语义信息与原始项目表示可能完全不同[22]。针对图结构数据的高阶传播知识相关性问题，基于异构传播的 HKIPN 模型提出的异构传播可以有效的提升传播范围。当前，随着传播层次增加，如何高效利用高阶交互信息是一大研究难点。本文受异质信息网络推荐传播思想[23]的启发，采用异构传播的方式通过分层知识感知注意力机制区分知识图谱实体邻居重要性，并增强用户、项目高阶表示以此解决高阶传播过程中的知识语义关联性问题。

2　推荐模型

本节介绍本文提出的基于用户潜在兴趣的知识感知传播推荐模型(Knowledge-aware propagation based on user's potential interest，KPUPI)，通过异构传播方法学习协同知识图中的高阶关联知识与用户兴趣表示。KPUPI 模型框架如图 1 所示。

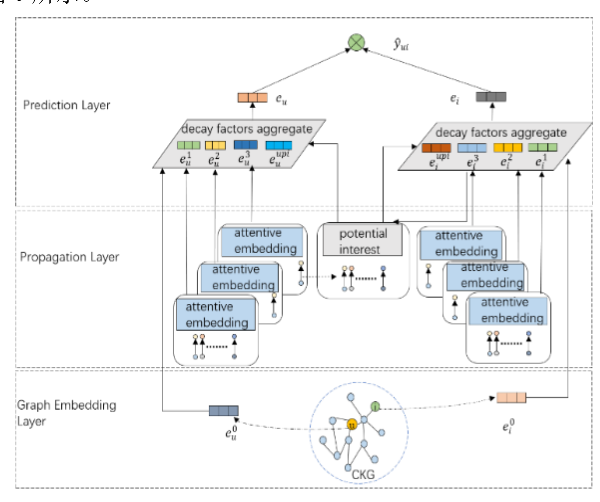


图 1　KPUPI 模型框架

Fig. 1　KPUPI model framework

模型主要分为 3 层，分别为图嵌入层、传播层以及预测层。图嵌入层生成协同知识图中的实体与关系的嵌入表示，嵌入表示中保留着图中的关联知识与图结构信息，e_u^0、e_i^0

分别表示用户与项目的初始化嵌入表示。在传播层中模型采用异构传播的方式同时传播关联知识与用户偏好，通过知识感知注意力机制区分相邻实体的重要性并获得用户与项目每层的嵌入表示 $\boldsymbol{e}_u^l$ 、 $\boldsymbol{e}_i^l$ ，在协同知识图上激活并学习用户高阶潜在兴趣以此增强用户最高阶的表示 $\boldsymbol{e}_u^{upi}$ 及项目最高阶的表示 $\boldsymbol{e}_i^{upi}$ 。预测层使用信息衰减因子区分不同传播层的信息衰减程度，并有效地聚合用户与项目的最终表示，通过内积操作得到预测分数 $\hat{y}_{ui}$ ，完成推荐预测。

2.1　图嵌入层

图嵌入层在三元组层次上得到关系与实体的嵌入表示，细粒度的学习具有结构信息的用户与项目表示，提升模型中用户和项目的表达能力，提高推荐性能。

知识图嵌入技术[24]是图表示学习领域的重要研究内容，该技术可以将协同知识图中的关联知识更好的表达出来，因此广泛应用于知识感知推荐算法中，本文使用 TransR 算法对协同知识图进行嵌入表示学习，该算法在不同的嵌入空间中建模实体与关系，不同的实体类型可以映射到同一关系空间中。给定协同知识图中的知识条目(h,r,t)，通过嵌入表示学习获得由关系空间映射得到的头实体与尾实体的嵌入表示 $\boldsymbol{h}_r$ 与 $\boldsymbol{t}_r$ ，定义如下：

$$\boldsymbol{h}_r = \boldsymbol{h}\boldsymbol{M}_r \,,\quad \boldsymbol{t}_r = \boldsymbol{t}\boldsymbol{M}_r \tag{1}$$

其中 $\boldsymbol{M}_r$ 为从实体空间映射到关系空间的转换矩阵，评分函数 $f_r(h,t)$ 计算两种实体对应关系空间中的偏差，其定义如下：

$$f_r(h,t) = \| \boldsymbol{h}_r + \boldsymbol{r} - \boldsymbol{t}_r \|_2^2 \tag{2}$$

其中 $\boldsymbol{h}_r$ 是头实体嵌入表示，$\boldsymbol{r}$ 为关系表示， $\boldsymbol{t}_r$ 是尾实体嵌入表示。评分函数 $f_r(h,t)$ 评分越小表明三元组为事实的概率更大。该算法的训练考虑到正例三元组与负例三元组之间的相对顺序，使用成对排名损失函数计算，定义如下：

$$L_{CKGE} = \sum_{(h,r,t,\tilde{t})\in T} -ln\sigma(f_r(h,t) - f_r(h,\tilde{t})) \tag{3}$$

其中 $T = \{(h,r,t,\tilde{t}) | (h,r,t)\in G, (h,r,\tilde{t})\notin G\}$ ，负例三元组是通过随机替换的方式生成的，σ(·)为 Sigmoid 激活函数。

2.2　异构传播层

2.2.1 知识感知传播

协同知识图中包含用户实体、项目实体以及项目的相关属性，实体之间的关系具有丰富的关联知识，并且，相邻实体之间具有强关联性，沿着协同知识图的链路传播可以获得逐层知识，随着传播层级增加可以得到不同层级的实体集与三元组，其中包含着高阶关联知识与用户兴趣。不同层次的有效信息可以丰富用户与项目的向量表示。

用户与项目每一层都存在大量的三元组集合，因此需要固定每层的三元组数量，用户与项目在图中的传播方式是相同的，本文以 o 为占位符，o 可表示用户 u 及项目 v，更高层次的三元组集合 E_o^l 中头实体 h 来源于上一层三元组集合 E_o^{l-1} ，通过层层传播的方式传递信息，其定义如下：

$$E_o^l = \{(h,r,t) | (h,t,t)\in G \ and \ h\in E_o^{l-1}\}, l=1,2,\ldots,L \tag{4}$$

其中 G 是知识图谱，l 为传播层次，在协同知识图上进行深度传播，可以有效的获取用户与项目的高阶交互信息，并提升用户与项目节点的表达能力。对一个用户或项目而言，不同关系对应的实体节点的重要性是不同的，因此在同一层的传播过程中，需要区分邻居的重要性。使用注意力机制可以有效地区分同一层次的传播过程中相邻实体之间的重要程度，使用户及项目的嵌入表示更加精确。

假设在协同知识图中的某用户或项目第 l 层的第 i 个三元组为(h,r,t)，加入知识感知注意力机制的尾实体节点嵌入表示 $\boldsymbol{a}_i$ 由尾实体嵌入表示 $\boldsymbol{e}_i^t$ 以及由注意力函数 $\pi(\cdot)$ 计算得到的对应注意力权重的乘积而来，定义如下：

$$a_i = \pi(e_o^0, e_i^h, r_i) e_i^t \tag{5}$$

$\boldsymbol{e}_o^0$ 为用户的初始表示， $\boldsymbol{e}_i^h$ 为头实体的嵌入表示， $\boldsymbol{r}_i$ 为关系实体的嵌入表示，注意力网络函数 $\pi(\cdot)$ 计算得到尾实体 $\boldsymbol{e}_i^t$ 的注意力权重，该权重可反映出头实体对尾实体节点的重要性，注意力函数定义如下：

$$z_0 = \mathrm{Re}LU(W_0(e_o^0 \| e_i^h \| r_i) + b_0) \tag{6}$$

$$\pi(e_o^0, e_i^h, r_i) = \sigma(W_2 \mathrm{Re}LU(W_1 z_0 + b_1) + b_2) \tag{7}$$

注意力网络通过拼接操作将实体初始表示 $\boldsymbol{e}_o^0$ 、头实体的嵌入表示 $\boldsymbol{e}_i^h$ 和关系实体的嵌入表示 $\boldsymbol{r}_i$ 拼接到一起，W 和 b 为需要学习的参数，不同的下标表示不同的层数，最后使用 Softmax 函数对三元组的系数进行归一化，注意力网络函数 $\pi(\cdot)$ 的最终定义如下：

$$\pi(e_o^0, e_i^h, r_i) = \frac{\exp(\pi(e_o^0, e_i^h, r_i))}{\sum_{(h',r')\in E_o^l} \exp(\pi(e_o^0, e_i^{h'}, r_i))_i} \tag{8}$$

E_o^l 为目标用户在第 l 层的三元组集合，聚合每层三元组中的尾实体的注意力嵌入表示 $\boldsymbol{a}_i$ 可以得到用户(项目)在传播第 l 层的嵌入表示 $\boldsymbol{C}_o^l$，定义如下：

$$\boldsymbol{C}_o^l = \sum_{i=1}^{|E_o^l|} \boldsymbol{a}_i \tag{9}$$

$|E_o^l|$ 为第 1 层中三元组的总数，每层的嵌入表示有效的表达传播过程中的细节，可以提高推荐模型的表达能力。

2.2.2 用户潜在兴趣传播

基于嵌入传播的知识图谱推荐在学习用户及实体的表示时在高阶传播过程中普遍存在语义相关性降低的问题。相较于 KGAT、HKIPN 等先进的推荐算法，本模型在高阶传播过程中采用用户潜在兴趣传播方式增强用户高阶表示，给定用户最高阶三元组集合 E_o^L ，通过比较用户交互的项目 v 与每个三元组中的头实体 $\boldsymbol{h}_i$ 与关系实体 $\boldsymbol{R}_i$ 分配一个相关性概率 p_i ，通过定义如下：

$$p_i = softmax(v^T \boldsymbol{R}_i \boldsymbol{h}_i) = \frac{exp(v^T \boldsymbol{R}_i \boldsymbol{h}_i)}{\sum_{(h,r,t)\in E_o^L} exp(v^T \boldsymbol{R}\boldsymbol{h})} \tag{10}$$

其中 $\boldsymbol{R}_i$ 与 $\boldsymbol{h}_i$ 分别是关系嵌入表示与头实体嵌入表示，可以通过相关性概率反映出在关联空间中项目与实体之间的相似性，其中不同关系条件中项目与对应实体的相似度不同。三元组中的尾实体 $\boldsymbol{t}_i$ 与对应的相关性概率 p_i 加权求和可获得用户的高阶潜在兴趣表示 $\boldsymbol{z}_u^L$，定义如下：

$$\boldsymbol{z}_u^L = \sum_{(h_i,r_i,t_i)\in E_o^L} p_i \boldsymbol{t}_i \tag{11}$$

用户高阶潜在兴趣表示可以有效的提高高阶传播中用户表示的表达能力。

2.3　预测层

在异构传播的过程中随着传播层级的增加存在关联知识相关性较小等信息衰减问题，通过神经网络检测项目在传播过程中产生的信息衰减信号，其衰减因子定义如下：

$$\boldsymbol{q}_0 = ReLU(W_3(\boldsymbol{e}_o^0 \| \boldsymbol{C}_o^l) + b_3) \tag{12}$$

$$s_o^l = \sigma(W_4 ReLU(W_4 \boldsymbol{q}_0 + b_4)) \tag{13}$$

其中使用 ReLu 函数作为非线性激活函数，σ(·)为 Sigmoid 函数，‖为拼接操作，W 和 b 为可学习的参数， s_o^l 为第 l 层的衰减因子。用户最终的表示 $\boldsymbol{e}_u$ 由用户初始表示 $\boldsymbol{w}_u^0$ 、用户在各层衰减因子影响下的层次表示以及用户高阶潜在兴趣表示 $\boldsymbol{z}_u^L$ 相结合，其定义如下：

$$\boldsymbol{e}_u = \boldsymbol{w}_u^0 + \sum_{i=1}^{L} s_u^l \boldsymbol{C}_u^l + \boldsymbol{z}_u^L \tag{14}$$

其中 $\boldsymbol{C}_u^l$ 为用户在第 l 层的嵌入表示。项目最终表示 $\boldsymbol{e}_v$ 由项目初始化表示 $\boldsymbol{w}_v^0$ 、项目在各层衰减因子影响下的层次表示 $\boldsymbol{C}_v^l$ 以及基于用户潜在兴趣的高阶项目表示 $\boldsymbol{z}_u^L$ 相结合，定义如下：

$$\boldsymbol{e}_v = \boldsymbol{w}_v^0 + \sum_{i=1}^{L} s_v^l \boldsymbol{C}_v^l + \Phi（\boldsymbol{z}_u^L + \boldsymbol{C}_v^L） \tag{15}$$

$\Phi(\cdot)$ 为全连接操作， $\boldsymbol{z}_v^L$ 为项目高阶交互表示， $\boldsymbol{C}_v^l$ 为项

目在第 l 层的嵌入表示。e_u、e_v 分别为用户与项目的最终表示，采用内积的方式预测用户与项目的偏好得分，得分大小代表用户对项目的交互概率：

$$\hat{y}_{uv} = e_u{}^{\top} e_v \tag{16}$$

2.4 损失函数

针对模型的优化问题，为了确保准确的实验效果，从用户未观测到的交互中随机抽取负交互，使正交互与负交互的样本大小相同。使用交叉熵损失函数评估推荐模型的效果，损失函数定义如下：

$$L_{CF} = \sum_{u \in U}\left(\sum_{v|(u,v)\in p^+} \Gamma(y_{uv}, \hat{y}_{uv}) - \sum_{v|(u,v)\in p^-} \Gamma(y_{uv}, \hat{y}_{uv})\right) \tag{17}$$

其中，Γ 为交叉熵损失函数，p^+ 表示用户的正交互，p^- 表示是用户的负交互。模型的最终损失函数包含协同知识图嵌入部分损失(公式(3))与推荐部分损失(公式(17))，并使用可学习的参数平衡两种损失值的重要性，使用 L_2 正则化防止过拟合，其定义如下:

$$L = L_{CF} + \lambda_1 L_{CKGE} + \lambda_2 \|\Theta\|_2^2 \tag{18}$$

其中 λ_1 为协同知识图嵌入损失与推荐部分损失的平衡超参数，$\Theta = \{\boldsymbol{U}, \boldsymbol{E}, \boldsymbol{R}, \boldsymbol{M}_r, \boldsymbol{W}_i, \boldsymbol{B}_i, \forall i \in \{0,1...4\}\}$ 是模型的一组参数，$\|\Theta\|_2^2$ 是由 λ_2 控制的 L_2 正则化项，最终对损失函数最小化学习模型的相关参数。

3 实验

本章节中使用两个真实场景下的公开数据集评估模型，并与当前最具有代表性的推荐算法进行对比分析，首先介绍实验所使用的数据集，对比的推荐算法以及实验参数设置，然后 CTR 点击率任务与 Top-K 推荐任务验证算法的推荐效果，通过实验获取模型最佳表现参数，最后通过消融实验证明模型中核心模块的重要性。

3.1 数据集介绍

为了有效地验证 KPUPIE 模型的有效性与实用性，选择两个真实场景下使用最广泛的知识图谱推荐公开数据集即音乐数据集 Last.FM[9]以及书籍数据集 Book-Crossing[10]进行实验，以上两个数据集都是以显示反馈的方式组织的，为了更好的反映模型推荐性能需要转换为隐式反馈。本文将数据集按照 6:2:2 的比例随机分为训练集、验证集和测试集，关于数据集的细节如表 1 所示。

表 1　实验使用的数据集

Tab. 1　The dataset used by the experiment

数据集	Last.FM	Book-Crossing
用户数	1872	17860
项目数	3846	14967
交互数	42346	139746
实体数	9366	77903
关系数	60	25
三元组数	15518	151500

3.2 对比方法与参数设置

为了进一步的实验验证模型的有效性，文章与以下经典的模型进行对比：

CKE[14]：一种基于嵌入的经典模型，将协同过滤与贝叶斯框架中的项目辅助信息结合，同时使用经典的 TransR 算法将知识图谱嵌入以此获取相关知识信息。

PER[25]：一种具有代表性的基于路径的推荐方法，使用元路径作为用户与项目之间的链接，提取基于元路径的实体特征。

RippleNet[18]：最经典的分层传播算法，以“涟漪”形式传播，获得用户的层次兴趣，增强用户表示。

KGCN[19]：基于非光谱图卷积网络的传播模型，通过区分邻居的重要性，有选择性的聚合邻居信息，有效的学习知识关联信息与用户的潜在兴趣以生成用户与项目的表示。

KGAT[9]：使用注意力机制在协同知识图谱中区分邻域信息的重要性，有效的学习高阶关系表示。

HKIPN[10]：使用异构传播方法，在协同知识图中分层传播，基于注意力机制生成不同权重的高阶邻居信息，以获得用户项目表示。

KPUPIE 模型的主要参数取值均参考主流的基于嵌入传播推荐模型中的实验参数范围进行对比，并在实验阶段分析数据集相关特征进行实验最终得出最优参数设置，其中学习率 Lr 在 $\{10^{-3}, 2\times10^{-3}, 10^{-2}, 2\times10^{-2}\}$ 中调节，L_2 正则化项的系数 λ 在 $\{10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}\}$ 中调节，嵌入向量维度 d 在 $\{8,16,32,64,128\}$ 中选择，层次传播深度 L 将在 $\{0,1,2,3,4\}$ 之间调整，层次三元组数 T 将在 $\{16,32,64,128\}$ 之间调整，使用 Adam 优化器对模型进行优化，并使用 Xavier 初始化器初始化模型参数，固定训练批次大小为 1024，通过实验获得最优参数设置如表 2 所示。最优参数的学习率 Lr 与系数 λ 均是默认参数，嵌入维度 d 通过分析数据集用户项目交互稀疏程度确定，为与对比方法传播效果进行直接对比，传播深度 L 及三元组数 T 均从较大数值中选取，突出模型特点。

表 2　实验结果最优参数

Tab. 2　The optimal parameters of the experimental results

数据集	Lr	λ	d	L	T
Last.FM	2×10^{-3}	10^{-5}	128	4	128
Book-Crossing	2×10^{-3}	10^{-5}	64	3	128

各对比方法实验主要参数的最佳设置都是通过实证研究或遵循原始论文来实现的，对比实验的所有对比方法均处于最佳状态下进行对比分析，保证对比实验有效性。

3.3 基线模型对比分析

KPUPIE 模型在 CTR 点击率预测任务过程中选择使用 F1 与 AUC 两个指标评估模型，在 Top-K 推荐任务中使用 Recall@K 评估模型。

AUC 指标反映模型对样本的排序能力，定义如下：

$$AUC = \frac{\sum_i^n (pos > neg) + 0.5 \times \sum_i^n (pos = neg)}{N_{pos} \times N_{neg}} \tag{19}$$

模型中 pos 表示正样本的得分，pos 表示负样本的得分；N_{pos}、N_{neg} 分别表示正样本、负样本的数量；n 表示样本总数量。F1 指标同时兼顾模型的准确率、精确率与召回率，是模型准确率与召回率的调和平均，定义如下：

$$F1 = \frac{2 \times P \times R}{P + R} \tag{20}$$

其中 P 为模型精确率,即被正确分类的正样本占被分为正样本的比例，R 为模型召回率,即所有正样本被正确找回的概率。在 TopK 推荐任务中选用的 Recall@K 评估模型，召回率反映用户推荐列表中出现测试集中项目的比例，定义如下：

$$\text{Recall}@K = \frac{\sum_{u \in U} |R(u) \cap T(u)|}{\sum_{u \in U} |T(u)|} \tag{21}$$

其中 R(u)为模型生成的用户 u 推荐列表，T(u)为用户 u 在测试集中的交互列表。

经过多次实验得到本模型在两个数据集上的 CTR 预测结果如表 3 所示，Top-K 推荐结果如图 2 所示，通过对实验结果的分析可以得到以下结论：

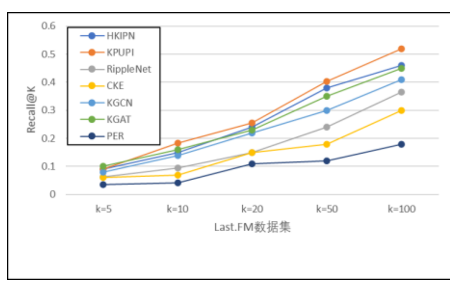
a) 本文提出的 KPUPIE 算法使用两个稀疏性较高的数据集，在 CTR 预测任务与 Top-K 推荐任务中都表现最好，具体而言，KPUPIE 模型在 Last.FM 数据集与 Book-Crossing

数据集上进行 CTR 预测任务，其中 AUC 指标相较于最先进的基线提升了 2.25%与 4.71%，F1 指标分别提升 3.05%和 1.20%。并且在 Top-K 推荐任务中 Recall@K 指标在绝大多数K值条件下都优于最先进的基线算法，充分的验证了本文提出的算法的有效性。

表 3 CTR 预测实验结果

Tab. 3 Result of CTR prediction experiment

模型	Last.FM		Book-Crossing	
	AUC	F1	AUC	F1
CKE	0.743	0.675	0.676	0.623
PER	0.641	0.601	0.605	0.573
RippleNet	0.778	0.702	0.721	0.647
KGCN	0.802	0.722	0.684	0.631
KGAT	0.825	0.739	0.731	0.654
HKIPN	0.845	0.754	0.743	0.666
KPUPI	**0.864**	**0.777**	**0.778**	**0.674**



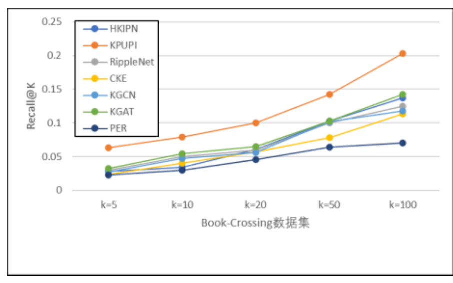


图 2 TOP-K 任务实验结果

Fig. 2 TOP-K task experimental results

b)对比基于知识图谱的推荐算法可知，在两个稀疏性较强的数据集上，基于传播的方法其推荐效果整体上优于基于嵌入的方法(CKE 模型)和基于路径的方法(PER 模型)。其原因在于基于传播的方法充分利用了实体的高阶连通性挖掘高阶关联信息，以此获取更好的用户和项目表示，使推荐结果更加精准。

c)与 KGCN 模型相比较，加入注意力机制的 KGAT 模型、HKIPN 模型和 KPUPIE 模型 AUC 值在两个数据集上至少提升 2.87%，说明结合注意力机制的推荐模型可以更精确的学习传播过程中用户和项目的嵌入表示。

d)RippleNet 模型、KGCN 模型、KGAT 模型、RippleNet 模型以及本文中提出的 KPUPIE 模型都是基于传播的方法，RippleNet 模型和 KGCN 模型从用户角度或项目角度单方面建模高阶连通性，并通过传播的方式挖掘用户兴趣(聚合项目的高阶邻居表示)，因此在学习用户-项目嵌入表示的过程中忽略部分有用信息。KGAT 模型和 HKIPN 模型的表现证明将用户-项目交互信息融合到知识图谱中可以同时有效地获取用户、项目的高阶信息，使推荐结果得到有效的提升。

e)与 RippleNet 模型从用户角度单方面的进行高阶传播表示学习相比较，本文模型的优势是在协同知识图上从用户与项目交互的角度进行高阶传播表示学习。与 KGAT 模型、HKIPN 模型相比较，本文模型进一步的改善高阶传播过程中的知识关联性，具体表现在异构传播层增加了用户与项目的高阶表示，更精确的学习用户的兴趣表示及项目表示。本文提出的 HPUPIE 模型在整体传播的过程中使用协同知识图，有效的学习传播过程中的用户、项目表示，并从用户角度利用高阶连通性增强用户高阶的表示，有效的解决高阶知识关联程度降低的问题，使推荐效果得到有效的提升。

3.4 实验参数影响分析

为了进一步的研究模型核心思想传播层异构传播对实验结果的影响，本节近一步的通过调节相关实验参数进行实验，分析不同网络结构对 CTR 预测任务中 AUC 指标的影响。

(1)KPUPIE 模型的嵌入维度设置如 3.2 小节所述，嵌入向量维度 d 在{8,16,32,64,128}中选择，实验结果如表 4 所示，通过实验分析，在一定程度上增加实体和关系的嵌入维度可以有效的提高推荐准确性。但 Book-Crossing 数据集实验结果显示，当维度太大时推荐准确性降低的情况，因为增加嵌入向量的维度可以编码更多的有效信息，但维度过高出现过平滑现象，致使推荐结果变差。

表 4 嵌入维度对 AUC 的影响

Tab. 4 The impact of embedding dimensions on AUC

维度	8	16	32	64	128
Last.FM	0.839	0.845	0.847	0.854	**0.864**
Book-Crossing	0.749	0.754	0.768	**0.778**	0.766

(2)模型中传播层数对推荐结果的影响从两个角度分析，增加传播层数可以有效的丰富用户与项目的表示，如表 5 所示，在 Last.FM 数据集上，增加传播层数使推荐结果的 AUC 指标明显的提高。但随着传播层数增加，越来越多的与用户、项目不相关的实体邻居被学习到用户及项目的表示中，产生了更多的噪声，一定程度上影响推荐性能。在 Book-Crossing 数据集上，AUC 指标在传播深度到达 4 层时比传播层数为 3 层更低。

表 5 传播层数对 AUC 的影响

Tab. 5 The influence of the number of propagation layers on AUC

层数	0	1	2	3	4
Last.FM	0.811	0.824	0.841	0.857	**0.864**
Book-Crossing	0.694	0.721	0.749	**0.778**	0.770

(1) 在传播过程中每个实体可能存在很多邻居，但由于计算开销的限定，文章仅讨论每层三元组数量在{16,32,64,128}不同大小时对推荐效果的影响。实验结果如表 6 所示，每层的三元组数在 128 时可以得到最好的推荐效果，由于 Book-Crossing 数据集和 Last.FM 数据集数据稀疏，需要更多的辅助信息提供关联知识丰富用户与项目的表示。

表 6 三元组数对 AUC 的影响

Tab. 6 The influence of the number of triples on AUC

三元组数	16	32	64	128
Last.FM	0.848	0.849	0.854	**0.864**
Book-Crossing	0.760	0.768	0.771	**0.778**

3.5 模型结构分析

本文以 CTR 预测任务的 AUC 值为消融实验的评估指标，讨论模型中衰减因子模块与用户最高阶兴趣增强模块在模型中有效性，KPUPI/D 模型不考虑传播过程中的信息衰减现象，将各层表示简单的聚合为最终的用户(项目)表示。KPUPI/U 模型在高阶信息传播学习用户(项目)的表示时不使用用户高阶兴趣增强模块，直接聚合相关表示。消融实验结果如表 7 所示，从表中可知：

a)当高阶信息传播过程中知识关联性降低问题存在时，在两个数据集上，KPUPI/U 模型的评估指标与 KPUPI 模型的表现相比有明显的差距，其原因是 KPUPI 模型使用用户潜在兴趣传播方法学习更多的用户、项目交互信息并且增加低阶实体与高阶实体的语义关联性。对比结果突显出信息传

播的过程中增强用户(项目)高阶嵌入表示的重要性。

b)KPUPI/D 模型简单的聚合各层的表示生成最终的表示并不能准确的表达用户及项目的最终表示，KPUPI 模型考虑信息传播的信息衰减情况，更精准的区分各层次的表示重要性，这表明在增加信息衰减因子，推荐效果在一定程度上会有所提升。

表 7　消融实验结果

Tab. 7　Results of ablation experiment

模型	KPUPI/D	KPUPI/U	KPUPI
Last.FM	0.860	0.8420	**0.864**
Book-Crossing	0.769	0.746	**0.778**

3.6　案例分析

本节针对用户兴趣传播进行案例分析，以展示用户高阶兴趣传播过程中具体的细节。为此，从 Last.FM 数据集中随机选取一个用户(u1670)构建与其交互相关的协同知识图示意图作为案例说明。如图 3 所示，该图描绘了用户 u1670 在协同知识图中进行三阶兴趣传播过程中的相关概率。图中可以观察到 KPUPI 模型区分项目知识图谱的项目实体相关属性对用户兴趣的不同影响。例如，用户的交互项目 i3623 在用户的一阶兴趣传播过程中相似性概率最高，说明该项目最能反映出用户的兴趣，因此可以为学习用户兴趣表示提供更多的信息，从而为用户兴趣在协同知识图中的高阶传播提供更精确的表示。

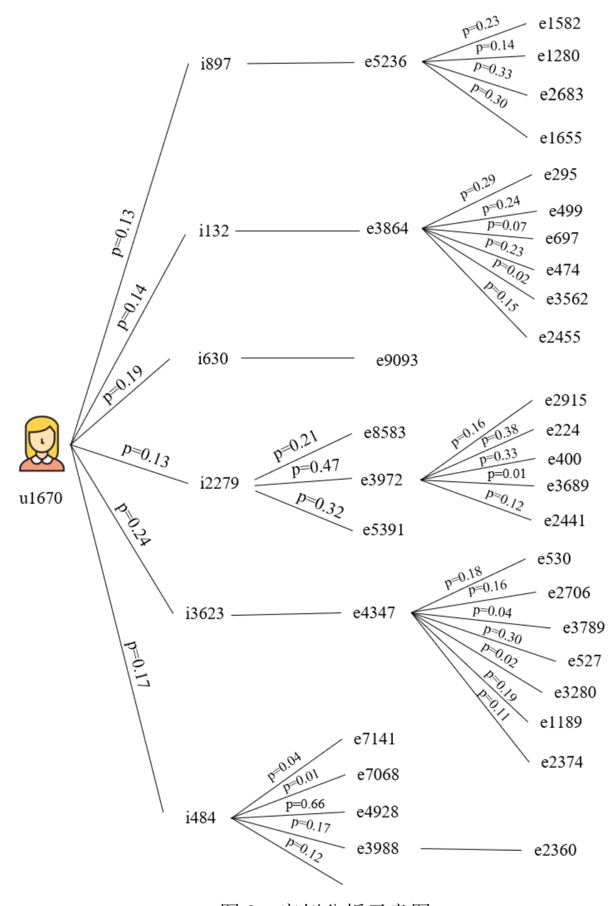


图 3　案例分析示意图

Fig. 3　Schematic diagram of a case study

4　结束语

本文提出基于用户潜在兴趣增强的知识感知传播推荐模型，在由用户-项目交互图与知识图谱相结合的协同知识图上异构传播，分层传播用户与项目的表示，在传播过程中使用知识感知注意力机制区分每层中邻居的重要性，有选择的聚合邻居节点表示，为解决高阶传播过程中存在信息衰减以及知识关联性减小的问题，通过在协同知识图的链接中自动迭代挖掘用户的高阶潜在兴趣，并补充用户最高阶表示，以此增强用户与项目的最终表示。使用音乐、书籍数据集与当前最经典的推荐模型进行实验对比，本算法的推荐效果显著提升。在未来，计划将用户社交网络结合到本文的模型中。将知识图谱与用户信息相融合，生成包含用户信息与项目信息的异质信息网络以此更好的建模用户与项目的交互关系、探索结构语义信息，增强推荐系统的准确性与可解释性。

参考文献：

[1] 赵俊逸，庄福振，敖翔，等. 协同过滤推荐系统综述 [J]. 信息安全学报, 2021, 6 (5): 17-34. (Zhao Junyi, Zhuang Fuzhen, Ao Xiang, *et al.* Survey of Collaborative Filtering Recommender Systems [J]. Journal of Cyber Security, 2021, 6 (5): 17-34.)

[2] Van Balen J, Goethals B. High-dimensional sparse embeddings for collaborative filtering [C]// Proc of the Web Conference 2021. New York: ACM Press, 2021: 575-581.

[3] Davis III K M, Spapé M, Ruotsalo T. Collaborative filtering with preferences inferred from brain signals [C]// Proc of the Web Conference 2021. New York: ACM Press, 2021: 602-611.

[4] Truong Q T, Salah A, Lauw H W. Bilateral variational autoencoder for collaborative filtering [C]// Proc of the 14th ACM International Conference on Web Search and Data Mining. New York: ACM Press, 2021: 292-300.

[5] Ji Shuyi, Feng Yifan, Ji Rongrong, *et al.* Dual channel hypergraph collaborative filtering [C]// Proc of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM Press, 2020: 2020-2029.

[6] Gope J, Jain S K. A survey on solving cold start problem in recommender systems [C]// 2017 International Conference on Computing, Communication and Automation (ICCCA) . Piscataway, NJ: IEEE Press, 2017: 133-138.

[7] 秦川，祝恒书，庄福振，郭庆宇，张琦，张乐，王超，陈恩红，熊辉. 基于知识图谱的推荐系统研究综述 [J]. 中国科学: 信息科学, 2020, 50 (07): 937-956. (Qin Chuan, Zhu Henshu, Zhuang Fuzhen, *et al.* Survey of Recommender S`ystem based on Knowledge Graph [J]. SCIENTIA SINICA Informationis, 2020, 50 (07): 937-956.

[8] Chicaiza J, Valdiviezo-Diaz. A Comprehensive Survey of Knowledge Graph-Based Recommender Systems: Technologies, Development, and Contributions [J]. Information (Switzerland) , 2021, 12 (6) .

[9] Wang Xiang, He Xiangnan, Cao Yixin, *et al.* KGAT: Knowledge graph attention network for recommendation [C]// Proc of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM Press, 2019: 950-958.

[10] Chen Qinghong, Tan Haobin, Lin Guangyan, *et al.* A Hierarchical Knowledge and Interest Propagation Network for Recommender Systems [C]// 2020 International Conference on Data Mining Workshops (ICDMW) . Piscataway, NJ: IEEE Press, 2020: 119-126.

[11] Jung Jaehun, Jung Jinhong, Kang U. Learning to walk across time for interpretable temporal knowledge graph completion [C]// Proc of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. New York: ACM Press, 2021: 786-795.

[12] Jia Z, Pramanik S, Saha Roy R, *et al.* Complex temporal question answering on knowledge graphs [C]// Proc of the 30th ACM International Conference on Information & Knowledge Management. New York: ACM Press, 2021: 792-802.

[13] Guo Liang, Yan Fu, Li Tian, *et al.* An automatic method for constructing machining process knowledge base from knowledge graph [J]. Robotics

and Computer-Integrated Manufacturing, 2022, 73: 102222.

[14] Zhang Fuzheng, Yuan N J, Lian Defu, *et al.* Collaborative knowledge base embedding for recommender systems [C]// Proc of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. New York: ACM Press, 2016: 353-362.

[15] Ji Guolaing, He Shizhu, Xu Liheng, *et al.* Knowledge graph embedding via dynamic mapping matrix [C]// Proceedings of the 53rd annual meeting of the association for computational linguistics and the 7th international joint conference on natural language processing. 2015: 687-696.

[16] Hu Binbin, Shi Chuan, Zhao Waynexin, *et al.* Leveraging meta-path based context for top-n recommendation with a neural co-attention model [C]// Proc of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM Press, 2018: 1531-1540.

[17] Wang Xiang, Wang Dingxian, Xu Canran, *et al.* Explainable reasoning over knowledge graphs for recommendation [C]// Proceedings of the AAAI conference on artificial intelligence. Palo Alto, CA: AAAI Press, 2019, 33 (01): 5329-5336.

[18] Wang Hongwei, Zhang Fuzheng, Wang Jiajin, *et al.* Ripplenet: Propagating user preferences on the knowledge graph for recommender systems [C]// Proc of the 27th ACM International Conference on Information and Knowledge Management. New York: ACM Press, 2018: 417-426.

[19] Wang Hongwei, Zhao Miao, Xie Xing, *et al.* Knowledge Graph Convolutional Networks for Recommender Systems [C]// In The World Wide Web Conference (WWW'19) . New York: ACM Press, 2019: 3307–3313

[20] Wang Hongwei, Zhang Fuzheng, Zhang Mengdi, *et al.* Knowledge-aware graph neural networks with label smoothness regularization for recommender systems [C]// Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining. New York: ACM Press, 2019: 968-977.

[21] Xu Zhuoming, Liu Hanlinh, Li Jian, *et al.* CKGAT: Collaborative Knowledge-Aware Graph Attention Network for Top-N Recommendation [J]. Applied Sciences, 2022, 12 (3): 1669.

[22] Wu S, Sun F, Zhang W, *et al.* Graph neural networks in recommender systems: a survey [J]. arXiv preprint arXiv: 2011. 02260, 2020.

[23] 刘佳玮, 石川, 杨成, 菲利普·俞. 基于异质信息网络的推荐系统研究综述 [J]. 信息安全学报, 2021, 6 (05): 1-16. (Liu Jiawei, Shi Chuan, Yang Cheng, *et al.* Heterogeneous Information Network based Recommender Systems: a survey [J]. Journal of Cyber Security) , 2021, 6 (05): 1-16.

[24] Wang Quan, Mao Zhengdong, Wang Bin, *et al.* Knowledge graph embedding: A survey of approaches and applications [J]. IEEE Trans on Knowledge and Data Engineering, 2017, 29 (12): 2724-2743.